COMPUTER SCIENCE TECHNICAL REPORT SERIES

UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND
20742

(15) DAAG53-76-C-Ø138,
WARPA Order-32Ø6

(11) 3Ø Apr 77

(6)

(12) 2Øp.

ALGORITHMS AND HARDWARE TECHNOLOGY FOR IMAGE RECOGNITION

(9) ~~Second~~ Semi-Annual Report no. 2,
1 November 1976 — 30 April 1977,

Contract DAAG53-76C-0138
(DARPA Order 3206)

Computer Science Center
University of Maryland
College Park, MD 20742

(10) Azriel Rosenfeld
David L. Milgram

ABSTRACT

Image models have been developed to propose thresholds, classify pixels and predict operator response to both signal and noise. A comprehensive approach to target cueing in FLIR imagery has been developed, based on simple but powerful heuristics. "Superslice", the algorithm embodying this approach, has been tested on two data bases and has successfully extracted object regions while accepting few noise regions. The target regions are discriminable from noise; however, they are difficult to identify reliably as to type. Westinghouse has designed a CCD implementation of a major portion of the Superslice algorithm.

The support of the U. S. Army Night Vision Laboratory under Contract DAAG53-76C-0138 (ARPA Order 3206) is gratefully acknowledged, as is the help of Mrs. Shelly Rowe.

403 018

1B

Foreword

This second semi-annual report summarizes the current status of the research being conducted under Contract DAAG53-76C-0138 (DARPA Order 3206), as well as plans for work to be done on this project in the near future. This project was initiated on May 1, 1976. It is being carried out by the Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD; Profs. Azriel Rosenfeld and David L. Milgram are principal investigators. It is devoted to the development and selection of algorithms for automatic target cueing on Forward-Looking InfraRed (FLIR) imagery, and to the hardware implementation of one or two such algorithms. The hardware aspects are being investigated by the Westinghouse Defense and Electronic Systems Center, Systems Development Division, Baltimore, MD; program director of this subcontract is Dr. Glenn E. Tisdale. The project is being monitored by Mr. John Dehne and Dr. George Jones of the U. S. Army Night Vision Laboratory, Ft. Belvoir, VA.

Certain portions of the technical material reviewed in this report have already appeared or will appear shortly as separate technical reports. This material will be only briefly described here.

Overall, the last six months (11/76-4/77) have seen the formulation of a comprehensive approach to the target cueing problem. This approach, based on the Superslice algorithm, makes use of powerful heuristics which appear to be well-founded for FLIR imagery: that object regions are thresholdable; that object regions can be distinguished from "accidents" by noting the relationship of the region interior to its edge and its surround; and that target regions are discriminable by shape and size from other object regions. Tests on two data bases substantiate both the ability of the approach to segment a FLIR scene into regions for further classification and its success in recognizing targets. The ability to identify the target regions as to type is still under development.

Westinghouse has studied the problem of implementing the Superslice algorithm and has designed circuitry for a number of modules. They have also considered the integration of the modules into a complete segmentation chip. The most crucial issue affecting the algorithm to date is the number and choice of thresholds to be used. Recognition performance demands a large number of thresholds, while physical constraints limit the number. This tradeoff is being actively studied.

1. Image Modelling

An approach to modelling FLIR imagery has been developed, based on the simplifying assumption that targets appear as homogeneous hot regions within a homogeneous cooler surround. This model describes the joint probability density of gray level and edge strength in such images, for various edge-detecting operators [1, 2]. In brief, the model predicts that for low edge values (corresponding to points in the interiors of objects and background), there should be two relatively well separated probability peaks, of different sizes, representing the gray levels of object and background interiors, respectively. For higher edge values, corresponding to points on object/background borders, these peaks should move together and become a single peak representing the border range of gray levels.

The model just described can be used as a guide to segmenting FLIR images by thresholding. At low edge values, it should be easy to pick a threshold at a gray level in the valley between the two probability peaks, since these are relatively well separated. At high edge values, the peak gray level value itself, or perhaps the mean gray level, should be a good threshold, since this represents the "center" of the edges. For intermediate edge values, one can compromise between these two thresholds in various ways. A comparative study of threshold selection schemes based on this approach has been conducted [3], and has shown them to be superior to conventional threshold selection models.

A different approach to modelling based on stochastic processes has also been developed [4]. Texture is represented as a general two-dimensional random Gaussian surface. The analysis is based on the two-dimensional energy spectrum. In the special case of isotropic texture, a bilateral autoregression model is suggested as being superior to conventional unilateral autoregression as used in time-series analysis.

Two studies dealing with the response of image operations to noise or background regions are in progress. One of them [5] presents a statistical analysis of the response of linear and non-linear image operators, including the Laplacian and several edge detectors. The input image is treated as a stationary random field of a context independent ensemble. The results are applicable to both FLIR and conventional imagery.

The second study is investigating the results of thresholding noise. The purpose of this study is to predict and control the false alarm rate associated with the Superslice algorithm (Section 3). Segmentation schemes based on thresholding must deal with accidental components which arise when a threshold "breaks up" a region. Knowledge of the expected size and shape of such regions can be used to build a classifier which can distinguish accidents from object regions. Working with synthetic noise images which have been smoothed in varying amounts, we have collected statistics on the number, size and shape of above threshold regions as a function of point probability.

As is expected, greater amounts of smoothing result in fewer but larger above-threshold components. Such components appear to be two-dimensional concatenations of thin line segments. Real FLIR images thresholded at background modes exhibit similar response curves, but we cannot as yet predict the component statistics from simple features such as probability or autocorrelation. An attempt is being made to devise a synthetic model with the same autocorrelation behavior as FLIR background regions. A complete report of the work will be written at a later date.

2. Object Extraction Based on Threshold Selection

As a preliminary to using thresholding to extract objects from an image, it is important to smooth the image, so that the extracted objects will not be too noisy. The use of both mean and median filtering for this purpose was investigated [1-2]. It was found that median filtering using a 5x5 neighborhood of each point produced the best results. An adaptive technique, which identifies neighborhoods that are noisy and edge-free, was shown in [6] to smooth noisy regions in images without degrading edges. The technique was also used to produce a weighting function to suppress spurious responses to an edge detector operating in a noisy environment.

Threshold selection based on the (gray level, edge strength) probability density has been investigated [1,2]. It was found that the average gray level of high gradient values was a good predictor of an object/background threshold. However, this selection scheme could not predict the multiple thresholds needed to segment multi-object scenes. In the following paragraph, we describe a new approach to threshold selection which provides a sequence of thresholds ordered by a figure of merit. There is no guarantee that every threshold will elicit object regions; however, preliminary investigations show that good thresholds are chosen by the algorithm.

We base our method on the heuristic "A good threshold segments object regions at edge points". Thus we choose thresholds which account for edge points. An alternative explanation is that the edge points of different object classes will cluster in different regions of (gray level, gradient value) space. Previous analysis has showed that raw edge detector response did not display local clustering behavior in the 2-D feature space. However, better behavior can be produced by first thinning the edge detector response using non-maximum suppression. Thinned edge points represent only a fraction of an image and do not cluster tightly, so that normal clustering methods are inappropriate. However, threshold selection can still take place, as will next be described.

Note that high gradient points associated with an object border can occupy a wider range of gray levels than similar low gradient points. This is because edge points are sampled from the ramp region of a profile, and the higher the ramp, the more gray levels are available. One therefore expects that clusters of high gradient points should vary more in gray level than clusters of low gradient points. Such clusters are likely to be cone shaped in either case. By counting the number of points in cone-shaped regions whose cusps lie on the grayscale axis and which increase in width with increasing gradient value, one can find that cone which contains the most edge points. After removing the contained points, the selection procedure can be reapplied iteratively. The algorithm terminates when an insignificant number of edge points remain or when every remaining cone picks up only an insignificant number of edge points. The cusp of each maximal cone selects a gray level which, as a threshold, "accounts" for the edge points within the cone.

The above algorithm has been modified to operate within disjoint gradient value ranges (as determined from the edge value histogram). This modification results in more predictable and desirable behavior. Naturally, not every selected threshold is a good one. On the other hand, good thresholds tend to be among those selected. It is hoped that this algorithm can be extended to other types of imagery.

3. Object Extraction Based on Edge/Border Coincidence

The approach which has been developed views the extraction of objects as a classification process into two classes: object regions and noise regions. Regions to be classified are extracted by first thresholding the (smoothed) image and then segmenting the thresholded image into connected components. Each connected component is considered to be a candidate for classification. Three heuristics are used: a size heuristic, a contrast heuristic, and a "well-definedness" heuristic. If object size range is known a priori, then noise regions outside the object size range can be rejected. The contrast heuristic states that objects contrast with their surrounds. This may be quantified by measuring the average gray level difference between the interior of a connected component and its boundary. Finally, the well-definedness heuristic states that objects are viewed as being distinct from their surround by the presence of an edge at the boundary. This is computed first, by extracting an "edge map" from the scene, consisting of the result of thinning the output of an edge detector; second, by measuring for each extracted region the percentage of its border which coincides with the edge map.

The combination of the contrast measure with edge/border coincidence serves both as a discriminant function for object regions and as a figure of merit for ranking the classified object regions. This approach does not require the user to preselect a particular threshold or set of thresholds. However, the speed of the algorithm is linear in the number of thresholds investigated. Moreover, the false alarm rate is related to the gray level probability of the chosen thresholds. This implies that care in selecting thresholds will generally be worthwhile (cf. Section 2). An implementation of this method ("Superslice") has provided good segmentations of FLIR windows. A more complete discussion is available in [7].

4. Region Tracking Using Dynamic Programming

In tracking object regions from frame to frame one can utilize the heuristic that region descriptions of the same object tend to cluster more closely than do descriptions of different objects. A method of tracking objects using dynamic programming to minimize a frame to frame discrepancy function has been successful on the small sequence data base which is currently available. The algorithm and test results are described in [8].

5. Target Classification

Regions classified as objects by the methods of Section 3 may be further classified as to target type. A hierarchical decision structure has been implemented, based on size, shape and contrast features. Object regions which survive the prescreening are divided into two groups based on size. The group of smaller regions is classified into target and noise classes based on compact shape and contrast. No attempt is made to identify the particular target types since these objects generally correspond to vehicles at long range with no identifiable characteristics. The group of larger regions is classified into tank, APC, truck and noise classes based on shape (compactness, symmetry, aspect) and contrast.

Selection of the set of features actually used at each node of the decision tree is restricted to those "logically allowable" at the given node. For example, while the brightness of a region is allowed to distinguish objects from noise, it is not used to determine vehicle type. The point of this restriction is to reduce the dependence of the final classifier on the pre-classified data, increasing both the robustness and the intelligibility of the classification. After this logical preselection is made, the effectiveness of features to be assigned to a node can be evaluated by standard statistical techniques (analysis of covariance, multiple discriminant analysis). The purpose is to increase the stability of the classifier without decreasing its accuracy. Experiments on the NVL data base exhibit good self-classification; however, we have not obtained as good results when extending the classifier to a test set. A discussion of these experiments will be forthcoming in the next quarter.

6. Image Processing Software

In an effort to clarify the implementation of the Superslice algorithm, a report [9] was written treating the question of data structures for region description. The algorithms described in the report build a tree structure for a thresholded image in which nodes correspond to connected components of object or background and in which the parent relation is based on region enclosure. In addition, the algorithm labels and computes features for each region and associates with the feature vector a chain encoding of the outer boundary of the region.

Another effort is currently in progress to design and build a generalized, concatenative local neighborhood image processing coordinator, called VIEWMASTER. The system is based on the premise that much image processing consists of a sequence of transformations defined on local neighborhoods of each point in which the output of each transform is input to the next. The conventional approach is to create intermediate image products corresponding to the output of each transform. This makes for very inefficient processing since much machine time and mass storage space is devoted to these unwanted images.

VIEWMASTER is a scheduler and supervisor of local neighborhood transformations. Intermediate images are avoided by devoting extra core storage to buffer the rows of transformed images. The generalized structure of the processing environment is quite similar to a PERT network with the attendant problems of resource allocation, race conditions, and deadlock. The successful implementation of this project will contribute to the efficient construction of large image processing programs.

7. Hardware Design

The Westinghouse Systems Development Division as a subcontractor to the University of Maryland has concentrated on the hardware implementation and fabrication of image algorithms for the focal plane [1, 2]. Algorithms whose hardware implementation has been designed include: median filtering, edge detection using differences of averages, edge thinning by non-maximum suppression, threshold selection based on a (gradient, gray level) histogram, noise cleaning by shrinking and expanding, and additional support logic such as serpentine delay lines and A/D converters. The attempt throughout is to design and build algorithms in analog CCD hardware within overall system constraints on data flow, storage requirements, chip size, yield factors and cost.

In recent quarterly reports [10, 11], Westinghouse has designed and analyzed the fabrication of CCD circuits for most of the Superslice algorithm. In particular, they have concentrated on the connected components module. Although not all issues have been resolved, substantial progress has been made in implementing this "intelligent" module in CCD.

8. Future Plans

Previous sections have discussed completed work and work in progress. The thrust of the continuing work and that which is beginning is the validation and extension of current techniques and a widening of the scope of investigation to include new varieties of image knowledge. The validation efforts include extensive data base testing, continued classifier development, further work on modelling, and investigation into the sensitivity of the algorithms to parameter change. Close interaction with Westinghouse will continue in the area of algorithm-hardware relationships.

We also plan to study the generality of the Superslice algorithm. Accordingly, we are extending it to extract regions in other types of imagery. This also supports a longer range goal of understanding current scenes from previous views of the scene. It is felt that the regions proposed by Superslice in the previous scene can provide a plan for understanding the current scene, by specifying what we can expect to extract.

Another area for future exploration is the use of additional knowledge sources, including range strobes, multispectral and temporal data, and situation context. Ultimately, one could develop a comprehensive model of the battlefield into which each type of knowledge can be integrated. An overall approach worthy of consideration is that of "common-sense algorithms" - a formal structure for representing tendencies and cause/effect relationships.

References

[1] Algorithms and Hardware Technology for Image Recognition, First Quarterly Report, Computer Science Center, Univ. of Maryland, College Pk., MD, July 1976.

[2] Algorithms and Hardware Technology for Image Recognition, First Semi-Annual Report, Computer Science Center, Univ. of Maryland, College Pk., MD, October 1976.

[3] D. P. Panda, Segmentation of FLIR Images by Pixel Classification, Computer Science Center, Univ. of Maryland, College Pk., MD, Tech. Rep. 508, Feb. 1977.

[4] D. P. Panda, Random Surface Representation of Texture, Computer Science Center, Technical Report (in preparation), Univ. of Maryland, College Pk., MD.

[5] D. P. Panda, Application of Some Statistical Results in Image Analysis, Computer Science Center, Technical Report (in preparation), Univ. of Maryland, College Pk., MD.

[6] D. P. Panda, A Method of Adaptive Smoothing and Edge Enhancement, Computer Science Center, Univ. of Maryland, College Pk., MD, Tech. Rep. 504, Feb. 1977.

[7] D. L. Milgram, Region Extraction Using Convergent Evidence, Proceedings of the ARPA SemiAnnual Workshop on Image Understanding, Minneapolis, MN, April 1977.

[8] D. L. Milgram, Region Tracking Using Dynamic Programming, Computer Science Center, Univ. of Maryland, College Pk., MD, Tech. Rep. 539, June 1977.

[9] D. L. Milgram, Constructing Trees for Region Description, Computer Science Center, Univ. of Maryland, College Pk., MD, Tech. Rep. 541, June 1977.

[10] A Discussion of Hardware Implementation and Fabrication for an Automatic Target Cueing System, Third Quarterly Status Report, Westinghouse, Systems Development Division, Feb. 1977.

[11] A Discussion of Hardware Implementation and Fabrication for an Automatic Target Cueing System, Fourth Quarterly Status Report, Westinghouse, Systems Development Division, May 1977.